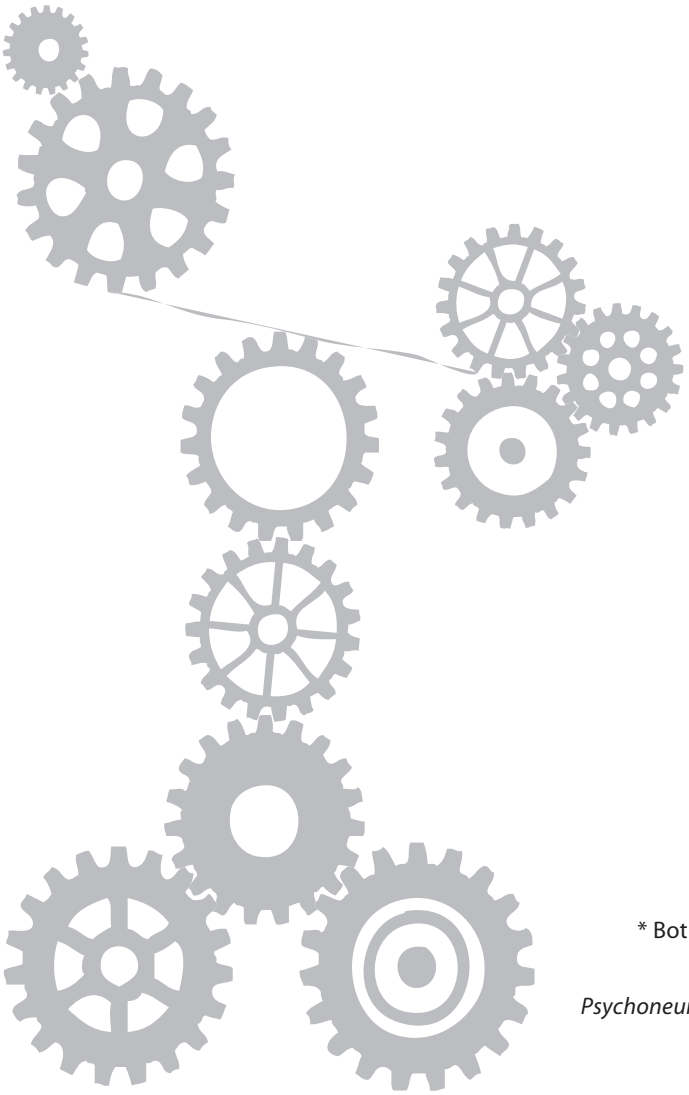
Chapter 2

The effect of two different doses of hydrocortisone on cognition in patients with secondary adrenal insufficiency – results from a randomized controlled trial



J Werumeus Buning*
P Brummelman*
J Koerts
RPF Dullaart
G van den Berg
MM van der Klauw
O Tucha
BHR Wolffenbuttel
AP van Beek

* Both authors contributed equally

Psychoneuroendocrinology 2015;55:36-47

ABSTRACT

Context

A wide variety in hydrocortisone (HC) substitution dose-regimens are considered physiological for patients with secondary adrenal insufficiency (SAI). However, it is likely that cognition is negatively influenced by higher cortisol exposure to the brain.

Objective

To examine the effects of a high physiological HC dose in comparison to a low physiological HC dose on cognition.

Design and Setting

This study was a randomized double blind cross-over study at the University Medical Center Groningen. This study is registered with ClinicalTrials.gov, number NCT01546922.

Patients

Forty-seven patients (29 males, 18 females; mean [SD] age, 51 [14] years, range 19–73) with SAI participated.

Intervention

Patients randomly received first a low dose of HC (0.2–0.3 mg/kg body weight/day) during 10 weeks followed by a high dose (0.4–0.6 mg/kg body weight/day) for another 10 weeks, or vice versa. HC substitution was given in three divided doses with the highest dose in the morning.

Main outcome measures

Cognitive performance (memory, attention, executive functioning and social cognition) of patients was measured at baseline and after each treatment period using a battery of 12 standardized cognitive tests.

Results

The higher dose of HC resulted in significantly higher systemic cortisol exposure for example measured at one hour after first dose ingestion (mean [SD], low dose: 653 [281] nmol/L; high dose: 930 [148] nmol/L; $P < 0.001$). No differences in cognitive performance were found between the two dose regimens.

Conclusions

No negative influence on memory, attention, executive functioning and social cognition was observed after 10 weeks of treatment with a higher physiological dose of HC in patients with SAI when compared to a lower dose.

INTRODUCTION

Patients with adrenal insufficiency are treated with glucocorticoids (GCs) to compensate for the loss of endogenous cortisol production. Usually this is done by oral administration of hydrocortisone (HC) or cortisonacetate (CA) with the aim to mimic the endogenous cortisol rhythm, with peak values in the early morning before waking and a nadir at bedtime. Body weight is an important determinator of the exposure of a dose measured using repeated pharmacokinetic sampling over a period of time.[1] Experts therefore recommend a weight-adjusted HC dose of 0.12 mg/kg body weight for the morning dose.[2] Furthermore, thrice-daily weight-adjusted administration mimics the day-curve of cortisol seen in healthy volunteers and is necessary because of the short half-life of HC.[1]

GC substitution therapy, although referred to as physiological, has its imperfections. Current GC dose-regimens inevitably result in over- or under-replacement during certain periods of the day which may result in poor quality of life.[2] Furthermore, Filipsson and colleagues found that GC substitution, especially in higher physiological doses (> 20 mg/day), was associated with an unfavorable metabolic profile when compared with patients with normal adrenal function.[3]

Besides the above mentioned side effects, cognitive side effects are reported in healthy individuals and in patients treated with pharmacological doses of GC. There is convincing evidence of an inverted "U"-shape relation between plasma cortisol levels and cognitive performance both in animals and humans.[4,5]

In healthy volunteers, an association has been found between higher cortisol levels and decreased memory performance,[5–12] and decreased executive functioning.[13] Furthermore, these associations are supplemented by findings from randomized controlled trials with healthy subjects further emphasizing the negative effect of elevated cortisol levels on memory[14,15] and executive functioning.[16] These impairments could be explained by the fact that memory and executive functioning rely on brain structures that contain a high concentration of GC receptors, such as the hippocampus[17] and the prefrontal cortex.[18] Indeed, Lupien and colleagues showed that chronically elevated cortisol levels are associated with reduced hippocampal volume and impairments in learning and memory tasks which depend upon the integrity of the hippocampus.[9] Furthermore, the prefrontal cortex is not only associated with executive functioning, but also with social cognition[19] and attention.[20] A study by Wolkowitz and colleagues showed that increased levels of cortisol may lead to less salient encoding of meaningful stimuli and may impair selective attention.[21] However, later studies suggest that cortisol does not affect attention.[14,15] Because attention is of great importance for almost all cognitive functions, it is important to further elucidate the effects of GCs on attention.

There seems to be compelling evidence of an association between levels of cortisol and several cognitive functions. However, most of the studies examining the effect of cortisol on cognition have been performed in healthy subjects, in which intact feedback regulation is likely to affect results. Furthermore, none of the studies has focused on the influence of administration for a substantial period of time of HC substitution on cognitive performance. Therefore, the aim of this study was to examine whether a low physiological HC substitution dose is better for cognition compared to a high physiological HC substitution dose, both administered for a substantial period of time, in patients treated for secondary adrenal insufficiency (SAI) who are characterized by absence of feedback regulation of the HPA-axis.

MATERIALS AND METHODS

Patients

In this randomized double blind cross-over study, patients were recruited for participation at the endocrine outpatient clinic of the University Medical Center Groningen (UMCG), a tertiary referral center for pituitary surgery in the Netherlands. A total of 63 patients were included in this study, of whom 60 patients completed the run-in phase and the baseline measurement (mean (SD) age, 52 (13), range 19–73, 35 males, 25 females). Eligibility, inclusion and follow-up is shown in Figure 1. All patients had SAI for which they received GC substitution therapy. To avoid effects of switching to a different type of GCs, all patients on CA were converted to treatment with HC in a bioequivalent dose, (i.e. CA dose (in mg) times 0.8 when compared to HC dose (mg)) during a four week run-in phase. The diagnosis of SAI was based on internationally accepted biochemical criteria, principally early morning (0800 – 0900 h) serum cortisol measurements and, if necessary, an insulin tolerance test. Early morning cut-off cortisol levels for adrenal insufficiency in our center were validated for patients with hypothalamic-pituitary disorders as previously published.[22] Thyroid hormone deficiency was based on a low serum free thyroxine concentration (< 11.0 pmol/l). Growth hormone (GH) deficiency was based on a low insulin-like growth factor 1 (IGF-1) Z-score (less than 2 SD) and/or an insufficient peak GH concentration (< 10 mU/l) in response to insulin-induced hypoglycaemia or a peak growth hormone < 18 mU/l during an arginine-GHRH test. Insufficiency of the pituitary-gonadal axis was defined in men as a testosterone concentration below 10 nmol/l, in premenopausal women (aged < 50 years) as loss of menses and in postmenopausal women (aged > 50 years) as LH and FSH concentrations below 15 mU/l. Diabetes insipidus was defined as the incapacity to properly concentrate urine (increased urine volume with low urine osmolality) in the face of a high plasma osmolality (and sodium) and/or current treatment with

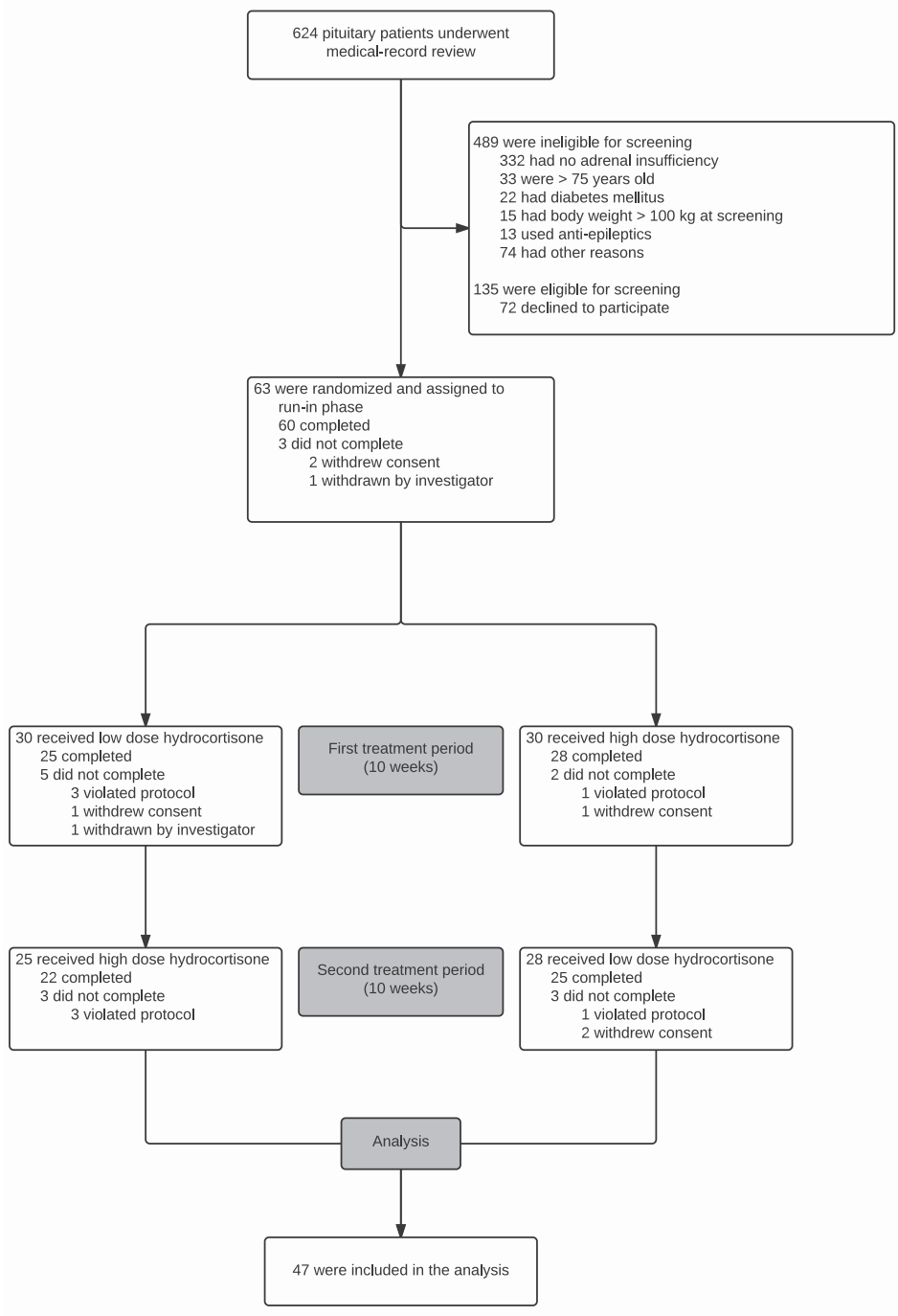


Figure 1. Eligibility, inclusion and follow-up of the patients.

desmopressin. Biochemical control of adequacy of hormonal substitution treatment was judged by the physicians that were responsible for the care of the participating patients using free thyroxine, IGF-1 and testosterone levels where necessary. Other inclusion criteria were age 18–75 years, body weight of 50–100 kg at screening, time interval of at least one year between study entry and tumor treatment with surgery and/or radiotherapy, and adequate replacement of all other pituitary hormone deficiencies for at least six months prior to entry of the study.

Main exclusion criteria were inability of legal consent, documented major cognitive impairment (MMSE < 24),[23] drug abuse or dependence, current psychiatric disorders, treatment for a malignancy, shift work, previous Cushing's disease, hospital admission during the study, diabetes mellitus with medication known to be able to induce hypoglycemia (e.g. Sulfonylurea derivatives and insulin) and a history of frequent episodes of clinical hypocortisolism. The concomitant use of other corticosteroids and drugs known to interfere with HC metabolism, e.g. anti-epileptics, was not allowed either.[2]

All patients were tested in the period between May 2012 and June 2013. The baseline characteristics of the entire cohort of patients with SAI (n = 135) known in our University Medical Center did not differ from the presented study population, confirming the representativeness of our study population (data not shown).

Ethics Statement

The study protocol was approved by the local ethics committee at the UMCG, The Netherlands. Patients gave written informed consent before entering the study. This study is registered with ClinicalTrials.gov, number NCT01546922.

Intervention

Patients were randomly assigned to either group 1 or group 2 by the research pharmacy with a block size of 4. Group 1 first received a physiological low dose of HC for 10 weeks, followed by a physiological high dose for another 10 weeks. Group 2 first received a physiological high dose for 10 weeks, followed by a physiological low dose (supplemental Fig. 1). Patients were treated with oral tablets containing either 5 mg HC (low dose) or 10 mg HC (high dose). Only the research pharmacy knew which dose was administered in each period. In the low dose condition, patients received a cumulative daily dose of 0.2–0.3 mg HC per kg body weight in three divided doses (before breakfast, before lunch, before diner). In the high dose condition, patients received the double amount, 0.4–0.6 mg HC per kg body weight. For the exact dosing schemes see Table 1. In cases of intercurrent illness or fever, patients were allowed to double or triple their HC dose. Because the study aimed to investigate two different dosing schemes, increasing the dose of HC was allowed for a maximum of 7 days (i.e. 10% of the study time and of the cumulative HC dose) but not in the week preceding

Table 1. Weight-adjusted dosages

Low dose				
Weight (kg)	**Before breakfast**	**Before lunch**	**Before diner**	**Total**
50–74	7.5	5.0	2.5	15
75–84	10.0	5.0	2.5	17.5
85–100	10.0	7.5	2.5	20
High dose				
Weight (kg)	**Before breakfast**	**Before lunch**	**Before diner**	**Total**
50–74	15.0	10.0	5.0	30
75–84	20.0	10.0	5.0	35
85–100	20.0	15.0	5.0	40

Dose is given in mg; Total: cumulative daily dose.

the second and the third visit. Compliance with the study medication was assessed in several ways. Firstly, by patient reports in daily diaries: patients were instructed to daily report if they had forgotten and/or doubled their medication and if so, how many doses they had forgotten or doubled. Secondly, the tablets returned by the patients after each study period were counted. Lastly, cortisol concentrations in plasma between the two study periods were compared.

Laboratory measurements

Serum cortisol was measured by a commercially available electrochemiluminescence immunoassay (ECLIA, Roche Modular Systems). Intra-assay coefficient of variations of serum cortisol were 1.5%, 1.1% and 0.9% at cortisol concentrations of 69.5, 348 and 952 nmol/l. The functional sensitivity was < 8.5 nmol/l with this assay. Cortisol in urine was analyzed by automated online solid phase extraction in combination with liquid chromatography tandem mass spectrometry (XLC-MS/MS) essentially as described by Jones et al..[24] The intra-assay coefficient of variation of urinary free cortisol was < 2.4% at a concentration of 57.2 nmol/l, and the inter-assay coefficient of variation was < 7.8% at a concentration of 54.3 nmol/l.

Cognitive tests

A battery of 12 standardized cognitive tests covering 4 cognitive domains (memory, attention, executive functioning and social cognition) was used. To assess memory function, the following tests were used: the Rivermead Behavioural Memory Test (RBMT), the 15 Words Test, the Digit Span Test (forward), the Rey Complex Figure Test and the 15 Figures Test. To assess attention, the subtests Vigilance, Divided Attention, Visual Scanning, and Alertness of the Test of Attentional Performance (TAP) were used. The

Verbal Fluency Tests, the Trail Making Test and the Digit Span Test (backwards) were used to evaluate executive functioning. Lastly, in order to assess social cognition the Reading the Mind in the Eyes Test was used. See supplemental materials and methods for a detailed description of the tests and reference data.

Questionnaire

A common questionnaire on demographic and health-related data was used to assess educational level, social status, full-time/part-time employment, social security benefit and use of medication. Educational level was classified using a Dutch education system, comparable to the International Standard Classification of Education (ISCED).[25] This scale ranges from 1 (elementary school not finished) to 7 (university level).

Procedure

At the three visit days (baseline [M1], 10 weeks later after the first treatment period [M2] and 10 weeks later after de second treatment period [M3]), patients were instructed to take their morning dose of HC at 07.00 h. At 08.00 h blood samples were drawn in a fasted state after which patients received breakfast. Breakfast was followed by a physical examination that included measuring body weight, height, and blood pressure. Between 08.30 h and 09.00 h the cognitive test battery started and took approximately 3 h including a 15 min break. During the break caffeine and nicotine intake was allowed. After finishing the cognitive test battery, the second blood samples were drawn (approximately 5 h after the morning HC intake). In addition, at the end of visit M1 and M2, study medication for the following treatment period and urine collection material was dispensed. During the day preceding M2 and M3, patients were instructed to collect 24 hour urine. All testing and scoring of tests was performed by trained personnel under the supervision of two neuropsychologists (PB and JK).

Safety

A total of 34 adverse events (AE) were reported in 29 of the 63 patients (46.0%). Two serious adverse events (SAE) were reported. One occurred on the low dose (hospitalization for influenza A infection) and one on the high dose (hospitalization for minor stroke in the left cerebral hemisphere). Both patients prematurely terminated the study due to this SAE as hospitalization was a criterion for withdrawal. No deaths occurred during the study.

A total of three patients withdrew from the study during the run-in phase, none of the withdrawals was related to HC (two withdrew their informed consent due to personal reasons, one stopped because of the inability to complete the test battery). A total of eight patients withdrew from the study while on a low dose, of which three withdrawals were possibly related to the HC dose (influenza A infection [n = 1], an inability to

tolerate the dose [n = 1] and a Herpes Zoster Ophthalmicus infection [n = 1]). Other reasons for withdrawal on the low dose were protocol violation (n = 3, incompliance with study medication [patient took an incorrect amount of tablets], initiation of anti epileptics and the use of prednisolone), broken arm (n = 1) and one patient withdrew the informed consent due to personal reasons. Five patients withdrew while on a high dose, of which one withdrawal was possibly related to the HC dose (unpleasant feelings). The other reasons for withdrawal were protocol violation (n = 3, incompliance with study medication (compliance was questioned by the investigator and suspicion was confirmed by the patient self and a relative), kenacort injection, changed job to shift worker) and a minor stroke in the left cerebral hemisphere (n = 1). The patients completing the study did not differ from those not completing the study with regard to age, sex, educational level, age at diagnosis, childhood or adult onset and body weight (data not shown).

With regard to the use of stress related dose adjustments, a doubling of the dose was reported 150 times in patients on the low dose and 146 times in patients on the high dose.

Statistical analyses

No relevant data can be inferred from literature to estimate a reasonable treatment effect of HC dose on cognition in secondary adrenal insufficiency. Because of the absence of relevant data from literature, we performed a power analysis. A study with 2 arms, each with 25 patients (total number of patients of 50) is able to detect an effect size of 0.4 (two-sided $\alpha = 0.05$ and $\beta = 0.80$) in test results even when between test correlations are poor (0.50). An effect size of 0.4 was chosen because it was considered a relevant change with a small to medium size effect. To allow for a drop-out rate of ±20% a total number of ±60 patients were needed. All statistical analyses were performed using Statistical Package for the Social Sciences (SPSS, Inc., Armonk, NY, USA), version 20. Demographic data are presented as median, interquartile ranges [IQR], frequencies or percentages. Cognitive performance data were presented as mean Z-score (SD). Higher Z-scores represent better cognitive performance compared to healthy subject of the same age, sex and educational level. Normality of data was analyzed using Q-Q plots. Since not all data were normally distributed, non-parametric tests for paired samples were used. To compare the cognitive performances at baseline of group 1 and group 2, the Mann-Whitney U-Test was used. The cognitive performance which was obtained while on a low dose of HC was compared to the performance on cognitive tests while on a high dose of HC by using the Wilcoxon Signed Rank Test. In addition, Cohen's d effect sizes were calculated to give a measure of the magnitude of the difference. An effect size of $d = 0.2$ is considered a small effect, $d = 0.5$ a moderate effect and $d = 0.8$ a large effect.[26] Furthermore, the performances of patients were compared to published normative datasets from healthy populations (see

supplemental materials and methods for a more detailed description), and standard scores were derived for each measure. A cognitive impairment on a test was defined as a performance equivalent to the performance of the lowest 10% of the reference samples.[23] To determine whether patients were more often impaired while on a high dose compared to a low dose, cognitive tests measuring similar cognitive processes were combined and the number of impaired patients on the high dose were compared to the number of impaired patients on the low dose using the McNemar Test. The two-tailed alpha level of < 0.05 was considered statistically significant. In case of statistical differences between the cognitive test performances on both doses, a Bonferroni correction was performed to correct for multiple comparisons. Given 42 test scores were compared, a *P*-value of < 0.001 (0.05/42) was regarded significant.

RESULTS

Study population

A total of 63 patients with SAI were included in this study. Forty-seven patients completed the study (29 men and 18 women, mean [SD] age 51 [14] years, range 19–73). Patients' characteristics are given in Table 2. Twenty-four patients received treatment for a pituitary adenoma (14 patients had a non-functioning macroadenoma, five patients had a macroprolactinoma and five patients had acromegaly). Three patients were treated for a cyst with pituitary localization, four were treated for a craniopharyngioma, five had a tumor distant from the pituitary gland (including one germinoma, one dysgerminoma, two meningeomas and one ENT tumor) and four had other acquired forms of ACTH deficiency (including traumatic brain injury). Furthermore, seven had a congenital form of ACTH deficiency (six combined, one isolated). Of the 32 patients who underwent surgery, 72% patients had undergone transsphenoidal surgery and 28% had undergone a craniotomy. The median [IQR] cumulative daily HC dose before the study was 25 [20; 30] mg. When corrected for kg body weight, the median [IQR] prestudy HC dose was 0.32 [0.25; 0.35] mg/kg body weight. Most patients (70%) received twice daily dosing prior to randomization. Group 1 and group 2 did not differ on clinical characteristics, except for the average time since radiotherapy in years (median [IQR], group 1, 20 [17; 31]; group 2, 10 [6; 18]; $P = 0.045$) and the number of daily doses of HC ([1/2/3], n; group 1, 2/11/9; group 2, 1/22/2; $P = 0.016$) (Table 2).

Serum and urinary cortisol

The administration of the morning dose of HC in the high dose condition (0.24 mg HC/kg body weight) resulted in significantly higher serum cortisol levels when compared to in the low dose condition (0.12 mg HC/kg body weight), both 1 hour and 5 hours

Table 2. Clinical characteristics of pituitary patients treated for secondary adrenal insufficiency ($N = 47$)

	Total patient group (N = 47)	Group 1 (N = 22)	Group 2 (N = 25)	P-value*
Age (y), median [IQR]	55 [43; 61]	55 [47; 61]	54 [40; 62]	0.966
Sex (males/females), *n*	29/18	13/9	16/9	0.730
Educational level (1/2/3/4/5/6/7)[a]	0/2/1/6/21/15/2	0/1/0/5/8/7/1	0/1/1/1/13/8/1	0.447
Age at diagnosis (y), median [IQR]	31 [20; 46]	31 [19; 42]	34 [21; 48]	0.316
Childhood onset/Adult onset of SAI, *n*	6/41	4/18	2/23	0.297
Body weight (kg), median [IQR]	82.5 [72.2; 93.0]	80.8 [71.2; 93.1]	85.4 [73.1; 93.2]	0.647
BMI (kg/m²), median [IQR]	26.6 [24.5; 29.4]	26.8 [24.7; 29.1]	25.9 [23.2; 30.3]	0.575
Surgery	32	13	19	0.215
Transsphenoïdal surgery/Craniotomy (%)	72/28	69/31	74/26	0.474
Age at surgery (y), median [IQR]	39 [28; 50]	38 [30; 42]	44 [20; 51]	0.502
Average time since surgery (y), median [IQR]	11 [6; 20]	18 [8; 29]	10 [4; 12]	0.074
Patients with 2nd surgery, *n*	5	0	5	0.052
Radiotherapy	19	6	13	0.085
Pituitary radiotherapy/cranial irradiation/radiotherapy for extracranial tumors (%)	84/11/5	83/0/17	85/15/0	0.213
Age at radiotherapy (y), median [IQR]	43 [25; 52]	32 [25; 40]	49 [24; 57]	0.171
Average time since radiotherapy (y), median [IQR]	12 [9; 22]	20 [17; 31]	10 [6; 18]	0.045
Hydrocortisone treatment prior to randomization				
Total daily dose (mg/day), median [IQR]	25 [20; 30]	23 [20; 30]	30 [20; 30]	0.186
Total dose/kg body weight/day (mg/kg/day), median [IQR]	0.32 [0.25; 0.35]	0.29 [0.25; 0.34]	0.34 [0.25; 0.37]	0.370
Number of daily dosings (1/2/3), *n*	3/33/11	2/11/9	1/22/2	0.014
Duration of hydrocortisone treatment (y), median [IQR]	12 [5; 22]	14 [7; 25]	10 [3; 14]	0.130

Table 2. Clinical characteristics of pituitary patients treated for secondary adrenal insufficiency ($N = 47$) (continued)

	Total patient group (N = 47)	Group 1 (N = 22)	Group 2 (N = 25)	P-value*
No. of hormonal replacements (1/2/3/4/5)	3/9/21/11/3	2/4/10/6/0	1/5/11/5/3	0.492
Thyroid hormone (% of patients substituted)	92	91	92	1.000
Growth hormone (% of patients substituted)	45	46	44	0.920
Growth hormone (% of patients unsubstituted)	21	28	35	1.000
Sex hormone (% of patients substituted)	57	46	68	0.119
Men: testosterone (% of patients substituted)	79	41	56	0.302
Premenopausal women, $n = 8$: estrogens (% of patients substituted)	50	33	60	1.000
Postmenopausal women, $n = 10$: estrogens	NA	NA	NA	
Desmopressin (% of patients substituted)	19	14	24	0.368

[a] Educational level was classified using a Dutch education system, comparable to the International Standard Classification of Education (ISCED).[25] This scale ranges from 1 (elementary school not finished) to 7 (university level). IQR: Interquartile range. SAI: Secondary adrenal insufficiency. NA: Not applicable.

* *P*-value of group 1 versus group 2.

Table 3. Serum cortisol and 24 h urinary free cortisol levels

	Low dose (N = 47)	High dose (N = 47)	P-value*
Serum cortisol levels			
1 hour after ingestion, nmol/L	653 (281)	930 (148)	< 0.001
5 hours after ingestion, nmol/L	208 (157)	312 (130)	< 0.001
Urinary free cortisol levels	(*N* = 46)	(*N* = 46)	
24 hour urinary cortisol, nmol/24 hr	88 (63)	339 (240)[a]	< 0.001

Data is given as mean (SD); *N* = 46 for urinary cortisol because of incontinence of one patient;
[a]*N* = 45 due to missing data;
* Low dose versus High dose by Wilcoxon.

after ingestion (Table 3). Furthermore, when comparing 24 hour urinary free cortisol levels, a high dose of HC resulted in significantly higher free cortisol levels compared to the low dose of HC (Table 3).

Cognitive tests

At baseline, group 1 (first receiving the low dose followed by the high dose) and group 2 (first receiving the high dose followed by the low dose) did not differ with regard to their performances on tests of cognition (supplemental Table 1). Cognitive performances of the 47 patients who completed the study are given in Table 4. The number of patients who showed an impaired performance on the cognitive tests are given in Table 5.

Memory performance

No significant differences between the doses were found for all memory measures (Table 4). Furthermore, there were no differences between both dose regimen in the number of patients scoring in the impaired range (Table 5).

Attention performance

No differences were found between the doses for neither the measures of attention (Table 4) nor the number of impaired patients (Table 5).

Executive functioning

Patients on a high dose did not differ from patients on a low dose on any of the measures of executive functioning (Table 4 and 5).

Table 4. Cognitive performances of pituitary patients treated for secondary adrenal insufficiency in the low and high dose condition

	Low dose (N = 47)	High dose (N = 47)	P-value*	High dose – Low dose	Effect size
Memory					
RBMT[d]					
Immediate memory	-0.41 (0.95)	-0.42 (1.33)	0.861	0.00 [0.20; 0.20]	0.01
Delayed memory	0.07 (0.98)	0.10 (1.33)	0.668	0.00 [0.20; 0.40]	0.02
Delayed corrected for immediate memory	0.85 (1.09)	0.89 (1.34)	0.713	0.00 [0.60; 0.70]	0.03
15 Words test[d]					
Short-term memory	0.36 (1.01)	-0.03 (1.03)	0.028[a]	-0.64 [-1.14; 0.00]	-0.38
Total immediate memory	0.43 (1.02)	0.25 (0.92)	0.227	-0.30 [-0.50; 0.20]	-0.19
Learning score	0.08 (1.05)	0.38 (1.19)	0.130	0.00 [0.00; 0.53]	0.26
Delayed memory	0.02 (1.04)	0.23 (0.97)	0.106	0.00 [-0.30; 0.40]	0.20
Delayed corrected for total memory	-0.30 (0.99)	-0.16 (0.99)	0.350	0.20 [0.00; 0.40]	0.14
Recognition	-0.11 (0.88)	0.08 (0.71)	0.093	0.00 [0.00; 0.00]	0.24
Digit Span forward[d]					
Short-term memory	0.49 (0.97)	0.56 (0.97)	0.614	0.00 [0.00; 0.36]	0.07
Rey Complex Figure[d]					
Immediate memory	1.14 (1.26)	1.40 (1.29)[b]	0.111	0.10 [-0.10; 0.50][b]	0.21
Delayed recall	1.11 (1.22)	1.28 (1.29)[b]	0.451	0.10 [-0.10; 0.40][b]	0.14
15 Figures Test (raw data)[d,f]					
Total memory	40.72 (14.64)	39.83 (14.18)	0.501	-0.98 [-3.38; 1.42]	-0.07
Delayed memory	10.02 (3.47)[c]	10.23 (3.80)	0.263	0.22 [-0.44; 0.87][c]	0.07
Recognition	28.36 (2.52)	28.43 (2.14)	0.606	0.04 [-0.69; 0.77]	0.02
Attention					
Vigilance (raw data)[d,f]					
Reaction time of correct responses	623.21 (85.65)	612.54 (117.88)[c]	0.385	-7.80 [-35.85; 20.24][c]	-0.10

Table 4. Cognitive performances of pituitary patients treated for secondary adrenal insufficiency in the low and high dose condition (continued)

	Low dose (N = 47)	High dose (N = 47)	P-value*	High dose – Low dose	Effect size
Variability of reaction time	122.02 (62.02)	130.39 (63.57)[c]	0.331	8.54 [-13.08; 30.17][c]	0.13
Number of omission errors	0.72 (2.41)	0.98 (2.47)[c]	0.732	0.28 [-0.18; 0.75][c]	0.04
Number of commission errors	1.81 (1.28)	1.91 (1.56)[c]	0.291	0.15 [-0.39; 0.69][c]	0.18
Divided attention[d]					
Reaction time auditory responses	-0.87 (0.81)	-0.89 (0.82)	0.905	0.00 [-0.28; 0.30]	0.03
Variability of reaction time auditory responses	0.03 (1.03)	-0.06 (1.00)	0.857	-0.09 [-0.50; 0.60]	-0.08
Reaction time visual responses	-0.06 (1.00)	0.03 (0.94)	0.502	0.00 [-0.20; 0.58]	0.09
Variability of reaction time visual responses	-0.14 (0.88)	0.07 (1.00)	0.277	0.08 [-0.17; 0.31]	0.23
Number of omission errors	-0.02 (0.86)	-0.14 (0.93)	0.531	0.00 [-0.51; 0.00]	0.14
Number of commission errors	0.30 (1.01)	0.34 (0.76)	0.681	0.00 [0.00; 0.40]	0.05
Visual scanning[e]					
Reaction time for target stimuli	-0.38 (1.04)	-0.36 (1.03)	0.713	0.00 [-0.27; 0.31]	0.01
Variability of reaction time for target stimuli	-0.48 (0.95)	-0.57 (0.92)	0.536	-0.11 [-0.30; 0.20]	-0.09
Number of omission errors	-0.03 (1.06)	0.08 (1.08)	0.229	0.10 [-0.11; 0.31]	0.10
Number of commission errors	-0.72 (0.26)	-0.70 (0.28)	0.329	0.00 [0.00; 0.00]	0.08
Alertness[e]					
Tonic alertness reaction time	-0.90 (0.73)	-0.99 (0.73)	0.192	-0.08 [-0.27; 0.00]	-0.13
Tonic alertness variability in reaction time	-0.33 (0.92)	-0.23 (1.04)	0.748	0.00 [-0.40; 0.37]	0.11
Phasic alertness reaction time	-1.03 (0.67)	-1.10 (0.66)	0.164	-0.17 [-0.28; 0.00]	-0.11
Phasic alertness variability in reaction time	-0.14 (0.82)	-0.42 (0.79)	0.015[a]	-0.38 [-0.51; 0.00]	-0.35
Executive functioning					
Fluency test[d]					
Semantic fluency	0.02 (1.24)	0.09 (1.18)	0.536	0.00 [-0.30; 0.60]	0.06
Phonemic fluency	0.32 (1.48)	0.21 (1.31)	0.572	-0.10 [-0.20; 0.40]	-0.08

Table 4. Cognitive performances of pituitary patients treated for secondary adrenal insufficiency in the low and high dose condition (continued)

	Low dose (N = 47)	High dose (N = 47)	P-value*	High dose – Low dose	Effect size
Digit Span backwards[d]					
Working memory	0.05 (0.86)	-0.01 (1.02)	0.632	0.00 [-0.31; 0.44]	0.07
Trail Making Test[e]					
Condition A, time	-0.32 (1.14)	-0.26 (1.06)	0.512	0.10 [-0.20; 0.40]	0.05
Condition A, no. of corrections (0/1)	43/4	43/4	1.000		
Condition B, time[g]	0.16 (1.46)	0.18 (1.31)	0.861	-0.30 [-1.30; 1.10]	-0.02
Condition B, no. of corrections (0/1/2/3)	34/7/4/1/1	30/9/6/1/1	0.403		
Condition B/A[g]	0.37 (1.42)	0.36 (1.48)	0.958	-0.60 [-1.40; 1.30]	-0.01
Social cognition					
Reading the Mind in the Eyes Test[d,f]	-0.67 (1.20)	-0.71 (1.15)	0.819	0.00 [-0.29; 0.29]	0.04

Cognitive performance data are given as mean Z-scores (SD) or as mean [95% confidence interval]; RBMT: Rivermead Behavioral Memory Test.

[a] *P*-value not statistically significant after Bonferonni correction.

[b] N = 45 (2 patients were excluded from analysis because of an impaired copy drawing, < 10th percentile) (See supplemental materials and methods).

[c] N = 46 due to a missing value.

[d] Verbal test.

[e] Non-verbal test.

[f] Cognitive performances on these tests are given as raw data mean (SD) due to lacking reference data.

[g] $P < 0.001$ for period effect.

* LD versus HD by Wilcoxon.

Table 5. Comparison of impaired scores between the low dose and the high dose of hydrocortisone

	Low dose (N = 47)	High dose (N = 47)	P-value*
Memory			
Immediate memory	13	15	0.727
Short-term memory	3	4	1.000
Delayed memory	8	8	1.000
Recognition	6	3	0.453
Attention			
Divided attention errors	6	7	1.000
Visual scanning errors	5	5	1.000
Executive functioning			
Fluency	10	10	1.000
Working memory	3	4	1.000
Cognitive flexibility	6	6	1.000
Social cognition	18	11	0.065
Psychomotor speed	17	24	0.092

Data is the number of patients showing an impaired score. Immediate memory: RBMT immediate memory + 15 Words Test total immediate memory + Rey's Complex Figure immediate memory; short-term memory: 15 Words Test short-term memory + Digit Span forward; delayed memory: RBMT delayed memory + 15 Words Test delayed memory + Rey's Complex Figure delayed memory; recognition: 15 Words Test recognition; divided attention errors: omission errors + commission errors; visual scanning errors: omission errors + commission errors; fluency: semantic fluency + phonemic fluency; working memory: Digit Span backwards; cognitive flexibility: Trail Making Test Condition B/A; social cognition: RMET; psychomotor speed: tonic alertness reaction time + phasic alertness reaction time + Trail Making Test Condition A.
* LD versus HD by McNemar Test.

Social cognition

With regard to social cognition no differences were found between the high and the low dose conditions (Table 4). Furthermore, there were no differences between the doses in the number of patients showing an impaired performance (Table 5).

DISCUSSION

This is the first randomized controlled double blind cross-over study demonstrating that there are no differences in the cognitive domains of memory, attention, executive functions and social cognition in patients with SAI treated with a physiological high dose of HC when compared to a physiological low dose of HC.

Although current substitution therapies for adrenal insufficiency are life-saving, optimization of treatment remains a challenge. Objective biochemical parameters to control dosage are lacking and the individual total dose is adjusted based on subjective parameters such as fatigue or quality of life.[27] A low dose of HC is recommended[3], but the clinical experience is that 40% of the patients receive a higher dose (> 30 mg HC equivalents/day).[3] The integrity of cognition is important for everyday functioning. Several cognitive functions are a basic requirement for participation in society, for example by maintaining a job and having a social life. Cognitive impairments can have an impact on these daily activities. The relationship between GC levels and cognitive performance has been studied in healthy subjects. However, no studies in patients with adrenal insufficiency were performed examining cognitive functioning in relation to HC dose.

The weight related dosing resulted in serum cortisol levels comparable with cortisol levels found by Mah et al..[1] The administration of a low dose of HC (0.12 mg HC/kg body weight for the morning dose) resulted in average serum cortisol levels of 646 ± 243 nmol/L 1 h after ingestion. In comparison, Mah found serum cortisol concentrations of 600–700 nmol/L 1 h after a single dose of 0.12 mg/kg body weight.[1] No significant differences were observed between the weight groups (50–74 kg, 75–84 kg, 85–100 kg) on the low dose and the high dose, both 1 h after ingestion and 5 h after ingestion (data not shown). The results of the current study thus confirm the reproducibility of weight related dosing. Furthermore, the higher dose of HC resulted in approximately one and a half times higher cortisol concentrations than the lower dose of HC.

The current study shows that there is no major negative effect of a higher dose of HC on memory performance. This finding is in accordance with a study by Newcomer and colleagues (1999) who did not find an effect of 4 days of treatment of 40 mg HC/day on memory performance in healthy subjects.[14] Furthermore, our results are in agreement with previous studies indicating that cortisol does not affect memory[28] and attention,[14,15] and executive functioning.[14] However, our findings are in contrast to studies in healthy volunteers showing an effect of higher levels of cortisol on memory performance[5–12,14,15] and executive functioning.[12] Considering social cognition, our findings are in contrast to a study showing a strong negative association between cortisol and social cognition in women.[29] Clear differences in study design and population may underlie this observed discrepancy as different doses and routes of administration were used. In addition, healthy volunteers have a circadian rhythm which is obviously lacking in SAI patients.

Without a Bonferroni correction for multiple testing, subtle differences were found between the low and high dose of HC regarding short-term memory and the variability of reaction time during the phasic alertness task, with patients on the high dose performing worse compared to patients on the low dose. This is in accordance with a

study performed by Harbeck et al., who found a negative correlation between cortisol levels and short-term memory in patients with primary adrenal insufficiency or SAI.[30]

The number of AE's, SAE's, and withdrawals were comparable on both dose regimens and between group 1 and group 2, indicating that there was no pattern in AE's and SAE's favoring one group or dose over the other. The AE's and SAE's reported can be considered a reflection of the events seen in daily practice in patients with SAI including more hospital admissions. Thus, in this view our patient sample can also be seen representative of the average population of patients with SAI.

Although this randomized controlled trial was performed according to the highest standards, some limitations need to be addressed. First, ten patients (21%) had IGF-1 Z-scores <-2 but did not receive GH substitution, indicating unsubstituted severe GH deficiency. This is usually the consequence of an absence of GH deficiency related symptoms or lack of patient benefit of previous GH treatment. Untreated GH deficiency may have influenced our results[31]. Even though all patients in this study received adequate replacement for all other pituitary hormone deficiencies prior to entry of the study and these substitution therapies remained unchanged during the study, other influences or interactions are possible and these might theoretically have also influenced our findings. Furthermore, some patients (40%) were previously treated with radiotherapy and some had a history of GH excess (11%). Previously, we have shown that neither relates to impaired cognitive functioning.[32,33] It was our goal to perform a study in a population representative of all patients with SAI visiting the endocrine outpatient clinic. In such a population, multiple covariates are present. Due to the cross-over study design, the influence of confounding covariates was reduced because each cross-over patient served as his or her own control. Secondly, learning effects were found for 2 tests (Trail Making Test condition B and Trail Making Test condition B/A). Thus a remarkably low period (learning) effect was found in our study design in which we used parallel version of most of the tests. A possible explanation for this learning effect of the Trail Making Test condition B and condition B/A was that patients developed a strategy after multiple testing, which may lead to the ameliorated performance. We also need to consider the possible influence of carry-over effects. For pharmacokinetic reasons, one would not expect this as it is found that blood concentrations of GCs reach peak levels after 15–30 min after the stressor, and decline slowly to pre-stress levels 60–90 min later.[34] Accordingly, Newcomer and colleagues showed that a difference on verbal memory performance between a higher dose of HC compared to a lower dose of HC and placebo were restored after a 6-day washout period.[14] However, cortisol is also known to induce genomic and structural changes in the limbic network.[35] Because of the vital function of cortisol, it is not possible to implement a wash-out period. Thus, even though we have corrected the best way we could for possible influences by our cross-over design, we cannot fully exclude learning and/or carry-over effects. Reassuring is that we did not find statistical evidence

for a carry-over effect. Thirdly, a heterogeneous patient group with regard to diagnosis and treatment was studied. However, we do believe that this is a realistic representation of the diagnoses seen in clinical practice. Therefore, our patient sample is representative of the patient group seen in practice. Lastly, one might argue that ten weeks is too short to find subtle differences. We believe however, that ten weeks is a reasonable time period to examine the effects of different doses of HC on cognition because some studies have found an effect of cortisol on cognition as early as after only 1 dose of GCs. Furthermore, other studies used similar time frames to assess clinically relevant endpoints.[36–38]

This study showed no differences in cognitive functioning between both doses in resting conditions. However, it has been found that glucocorticoids enhance memory consolidation of emotionally arousing experiences, but impair memory retrieval and working memory during emotionally arousing test situations.[39] It is therefore possible that emotional stress induction yields different effects on cognition in our patients.

In conclusion, we found no differences in cognition in patients with SAI on a physiological high dose of HC for 10 weeks when compared to a physiological low dose of HC. Therefore, both doses can be considered safe to administer with regard to cognitive functioning. However, opposite effects of higher doses of GCs on, for example, metabolic profile,[3] bone density,[40] and quality of life[37,41] need to be taken into consideration.

ACKNOWLEDGEMENT

We would like to thank Dr. W. J. Sluiter for his help with the statistical analyses.

REFERENCES

1. Mah PM, Jenkins RC, Rostami-Hodjegan A, et al. Weight-related dosing, timing and monitoring hydrocortisone replacement therapy in patients with adrenal insufficiency. *Clin Endocrinol (Oxf)* 2004;61(3):367-375.
2. Debono M, Price JN, Ross RJ. Novel strategies for hydrocortisone replacement. *Best Pract Res Clin Endocrinol Metab* 2009;23(2):221-232.
3. Filipsson H, Monson JP, Koltowska-Haggstrom M, Mattsson A, Johannsson G. The impact of glucocorticoid replacement regimens on metabolic outcome and comorbidity in hypopituitary patients. *J Clin Endocrinol Metab* 2006;91(10):3954-3961.
4. Gallagher P, Reid KS, Ferrier IN. Neuropsychological functioning in health and mood disorder: Modulation by glucocorticoids and their receptors. *Psychoneuroendocrinology* 2009;34 (suppl 1):S196-207.
5. Lupien SJ, McEwen BS. The acute effects of corticosteroids on cognition: integration of animal and human model studies. *Brain Res Brain Res Rev* 1997;24(1):1-27.
6. Kirschbaum C, Wolf OT, May M, Wippich W, Hellhammer DH. Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. *Life Sci* 1996;58(17):1475-1483.
7. Lupien S, Lecours AR, Lussier I, Schwartz G, Nair NP, Meaney MJ. Basal cortisol levels and cognitive deficits in human aging. *J Neurosci* 1994;14(5 Pt 1):2893-2903.
8. Lupien SJ, Wilkinson CW, Briere S, Ng Ying Kin NM, Meaney MJ, Nair NP. Acute modulation of aged human memory by pharmacological manipulation of glucocorticoids. *J Clin Endocrinol Metab* 2002;87(8):3798-3807.
9. Lupien SJ, Fiocco A, Wan N, et al. Stress hormones and human memory function across the lifespan. *Psychoneuroendocrinology* 2005;30(3):225-242.
10. Maheu FS, Collicutt P, Kornik R, Moszkowski R, Lupien SJ. The perfect time to be stressed: a differential modulation of human memory by stress applied in the morning or in the afternoon. *Prog Neuropsychopharmacol Biol Psychiatry* 2005;29(8):1281-1288.
11. Newcomer JW, Craft S, Hershey T, Askins K, Bardgett ME. Glucocorticoid-induced impairment in declarative memory performance in adult humans. *J Neurosci* 1994;14(4):2047-2053.
12. Young AH, Sahakian BJ, Robbins TW, Cowen PJ. The effects of chronic administration of hydrocortisone on cognitive function in normal male volunteers. *Psychopharmacology (Berl)* 1999;145(3):260-266.
13. McCormick CM, Lewis E, Somley B, Kahan TA. Individual differences in cortisol levels and performance on a test of executive function in men and women. *Physiol Behav* 2007;91(1):87-94.
14. Newcomer JW, Selke G, Melson AK, et al. Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Arch Gen Psychiatry* 1999;56(6):527-533.
15. de Quervain DJ, Roozendaal B, Nitsch RM, McGaugh JL, Hock C. Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nat Neurosci* 2000;3(4):313-314.
16. Lupien SJ, Gillin CJ, Hauger RL. Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose-response study in humans. *Behav Neurosci* 1999;113(3):420-430.
17. Joels M. Functional actions of corticosteroids in the hippocampus. *Eur J Pharmacol* 2008;583(2-3):312-321.

18. Sanchez MM, Young LJ, Plotsky PM, Insel TR. Distribution of corticosteroid receptors in the rhesus brain: relative absence of glucocorticoid receptors in the hippocampal formation. *J Neurosci* 2000;20(12):4657-4668.
19. Van Overwalle F. Social cognition and the brain: a meta-analysis. *Hum Brain Mapp* 2009;30(3):829-858.
20. Li C, Chen K, Han H, Chui D, Wu J. An FMRI study of the neural systems involved in visually cued auditory top-down spatial and temporal attention. *PLoS One* 2012;7(11):e49948.
21. Wolkowitz OM, Reus VI, Weingartner H, et al. Cognitive effects of corticosteroids. *Am J Psychiatry* 1990;147(10):1297-1303.
22. Dullaart RP, Pasterkamp SH, Beentjes JA, Sluiter WJ. Evaluation of adrenal function in patients with hypothalamic and pituitary disorders: comparison of serum cortisol, urinary free cortisol and the human-corticotrophin releasing hormone test with the insulin tolerance test. *Clin Endocrinol (Oxf)* 1999;50(4):465-471.
23. Lezak MD, Howieson D, Loring DW. Neuropsychological Assessment. 4th ed. New York: Oxford University Press; 2004.
24. Jones RL, Owen LJ, Adaway JE, Keevil BG. Simultaneous analysis of cortisol and cortisone in saliva using XLC-MS/MS for fully automated online solid phase extraction. *J Chromatogr B Analyt Technol Biomed Life Sci* 2012;881-882:42-48.
25. United Nations Educational, Scientific and Cultural Organization. International Standard Classification of Education (ISCED). 2006.
26. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
27. Arlt W, Rosenthal C, Hahner S, Allolio B. Quality of glucocorticoid replacement in adrenal insufficiency: clinical assessment vs. timed serum cortisol measurements. *Clin Endocrinol (Oxf)* 2006;64(4):384-389.
28. Alhaj HA, Massey AE, McAllister-Williams RH. Effects of cortisol on the laterality of the neural correlates of episodic memory. *J Psychiatr Res* 2008;42(12):971-981.
29. Smeets T, Dziobek I, Wolf OT. Social cognition under stress: differential effects of stress-induced cortisol elevations in healthy young men and women. *Horm Behav* 2009;55(4):507-513.
30. Harbeck B, Kropp P, Monig H. Effects of short-term nocturnal cortisol replacement on cognitive function and quality of life in patients with primary or secondary adrenal insufficiency: a pilot study. *Appl Psychophysiol Biofeedback* 2009;34(2):113-119.
31. Oertel H, Schneider HJ, Stalla GK, Holsboer F, Zihl J. The effect of growth hormone substitution on cognitive performance in adult patients with hypopituitarism. *Psychoneuroendocrinology* 2004;29(7):839-850.
32. Brummelman P, Elderson MF, Dullaart RP, et al. Cognitive functioning in patients treated for non-functioning pituitary macroadenoma and the effects of pituitary radiotherapy. *Clin Endocrinol (Oxf)* 2011;74(4):481-487.
33. Brummelman P, Koerts J, Dullaart RP, et al. Effects of previous growth hormone excess and current medical treatment for acromegaly on cognition. *Eur J Clin Invest* 2012;42(12):1317-1324.
34. de Kloet ER, Joels M, Holsboer F. Stress and the brain: from adaptation to disease. *Nat Rev Neurosci* 2005;6(6):463-475.
35. Joels M, Karst H, Krugers HJ, Lucassen PJ. Chronic stress: implications for neuronal morphology, function and neurogenesis. *Front Neuroendocrinol* 2007;28(2-3):72-96.

36. Johannsson G, Nilsson AG, Bergthorsdottir R, et al. Improved cortisol exposure-time profile and outcome in patients with adrenal insufficiency: a prospective randomized trial of a novel hydrocortisone dual-release formulation. *J Clin Endocrinol Metab* 2012;97(2):473-481.
37. Benson S, Neumann P, Unger N, et al. Effects of standard glucocorticoid replacement therapies on subjective well-being: a randomized, double-blind, crossover study in patients with secondary adrenal insufficiency. *Eur J Endocrinol* 2012;167(5):679-685.
38. Ekman B, Bachrach-Lindstrom M, Lindstrom T, Wahlberg J, Blomgren J, Arnqvist HJ. A randomized, double-blind, crossover study comparing two- and four-dose hydrocortisone regimen with regard to quality of life, cortisol and ACTH profiles in patients with primary adrenal insufficiency. *Clin Endocrinol (Oxf)* 2012;77(1):18-25.
39. de Quervain DJ, Aerni A, Schelling G, Roozendaal B. Glucocorticoids and the regulation of memory in health and disease. *Front Neuroendocrinol* 2009;30(3):358-370.
40. Lovas K, Gjesdal CG, Christensen M, et al. Glucocorticoid replacement therapy and pharmacogenetics in Addison's disease: effects on bone. *Eur J Endocrinol* 2009;160(6):993-1002.
41. Ragnarsson O, Mattsson AF, Monson JP, et al. The relationship between glucocorticoid replacement and quality of life in 2737 hypopituitary patients. *Eur J Endocrinol* 2014;171(5):571-579.

SUPPLEMENTAL DATA

Supplemental Materials and Methods. Full description of cognitive tests and reference data

The following cognitive domains were assessed using several standardized and validated cognitive tests: memory, attention, executive functioning and social cognition.

Memory

During the Rivermead Behavioral Memory Test (RBMT),[1] a short story was read out loud to patients who were then asked to immediately recall as many details as possible from the story. This allowed the calculation of an immediate memory score for coherent verbal information. A delayed memory score for coherent verbal memory was determined based on the number of details of the story recalled by patients after a period of about 30 minutes. Different stories were used during the second and third visit. Verbal memory was also assessed with the 15 Words Test, which is a Dutch equivalent of the Rey Auditory Verbal Learning Test.[2] During this test, 15 words were presented five times. After each trial, patients were asked to name immediately the words they remembered. This allowed the calculation of three different scores describing immediate memory. The short-term memory score is based on the number of words patients were able to name after the first presentation of the word list. The total immediate memory score represents the total number of words the patients remembered over the five trials. The learning score describes the difference between the number of words remembered in the third trial in comparison with the first trial. Besides immediate memory, delayed memory was measured with the 15 words test. The delayed memory score is based on the number of words a patient could recall after a period of about 30 minutes. To correct for the total immediate memory, there was also a delayed corrected for total memory score. The last part of the test was a recognition task. A total of 30 words were presented to the patient, of which 15 words were presented during the previous five trials and of which 15 words were new to the patient. The patients had to decide whether the words were new or familiar, which allowed the calculation of a recognition memory score. During the second and third visit, different lists of words were applied.

To assess short term memory, the Digit Span Forward test was used.[3] Patients were presented with a sequence of digits and immediately had to repeat the digits in the same order as they were presented. The test started with sequences of two digits which expanded until the maximum of nine digits. When a patient could not correctly repeat two sequences of the same length, the test was aborted. The total number of correctly repeated sequences was registered. During the second and third visit, different versions of the test were performed. Visual memory was assessed with the Rey Complex Figure

Test.[4] During this test, a complex geometrical figure was placed in front of patients, who were instructed to copy the figure as accurately as possible. After a 3-minute delay, patients were asked to reproduce the figure from memory. Approximately 15 minutes later, patients were again asked to reproduce the figure from memory. The number of details correctly recalled after 3 and 15 minutes were used to calculate an immediate and delayed memory score. If patients had an impaired copy drawing (< 10^{th} percentile) measures of immediate and delayed memory were excluded for analysis. An impaired copy drawing might indicate problems with visuoconstructive functions causing unreliable scores for immediate and delayed memory. During the second and third visit, different complex figures were used.

Aspects of visual memory were also assessed with the 15 Figures Test. During this test, 15 simple, geometrical figures were presented five times. After each trial, patients were asked to draw immediately the figures they remembered. This provides a score for immediate visual memory. The delayed visual memory score was based on the number of figures a patients could recall after a period of 15 minutes. Finally, a recognition task was administered, which consisted of 30 figures, of which 15 were presented during the five trials and of which 15 figures were new. Patients had to decide whether the figure was new of familiar, which allowed the calculations of a recognition score for visual memory. To prevent learning effects, different figures were used during the assessment after each treatment period.

Attention

The Test of Attentional Performance is a computerized test consisting of several subtests.[5] In the present study, the subtests vigilance, divided attention, visual scanning and alertness were used. Test instructions were presented on a computer screen and read out loud by the test examiner. In order to familiarize the participants with the tasks, a brief sequence of practice trials preceded each test.

During the vigilance task, two vertically arranged squares were shown in the middle of the screen and a pattern jumped from one square to the other. However, every now and then the pattern appeared twice in succession in the same square. When this happened, patients had to respond as quickly as possible by pressing a key. Reaction time of correct responses, variability in reaction time and the number of omission (lack of response to target stimuli) and commission (response to non-target stimuli) errors were calculated.

Divided attention is needed when two (or more) tasks have to be performed simultaneously. In the visual task, a series of matrices was presented in the center of the computer screen. Each matrix (4x4) consisted of a regular array of sixteen dots and crosses. When four of these crosses formed a square, patients had to press a key as quickly as possible. The acoustic task consisted of a high (2000 Hz) and low (1000

Hz) tone which were presented alternately according to the synchronous rhythm of the changing positions of the crosses. When the same tone occurred twice in a row, patients had to press a key as quickly as possible. A total of 100 visual and 200 acoustic stimuli was presented including 17 visual and 16 acoustic targets. Reaction time of correct responses, variability in reaction time and the number of omission and commission errors were calculated and represented measures of divided attention.

During the visual scanning task a series of 5 by 5 matrices were presented. Each matrix consisted of a regular array of 25 squares which each had an opening on one side (top, bottom, left or right side). A square with an opening on the top was defined as the target stimulus. The target stimulus occurred in 50% of matrices and was randomly distributed across the matrix. Patients were asked to press the left response button as quickly as possible when a matrix contained the target stimulus and to press the right response button if the target stimulus was not present. A total of 50 trials was presented (25 target trials and 25 non-target trials). Reaction times of correct responses, variability of reaction time and the number of commission and omission errors were calculated and represented measures of selective attention, i.e. attention for the target stimulus and not for non-target stimulus.

During the alertness task, reaction times were measured during two conditions. The first condition concerned a simple reaction time measurement, during which a cross appeared on the screen at random varying intervals. Patients were instructed to respond as quickly as possible by pressing a key when the cross appeared (tonic alertness task). In the second condition, a warning tone preceded the appearance of the stimulus (phasic alertness task). A total of 40 trials were presented, 20 trials with warning tone and 20 trials without warning tone. The time span between the warning tone and the appearance of the stimulus was random (between 300 and 700 ms). Measures of tonic and phasic alertness were calculated on the basis of reaction times of patients. In addition, the variability of reaction time and the number of omission errors were measured.

Executive functioning

Divergent thinking is the ability to approach a task or situation in different ways. So-called fluency tests rely on divergent thinking and evaluate the spontaneous production of words under restricted search conditions.[6] During the first fluency test applied in the present study, the semantic subtest, patients were asked to name as many animals as possible within one minute. For the second and third measurement patients were asked to name as many professions and grocery articles respectively. Participants were not allowed to name the same word twice. In a second fluency test, the phonemic subtest, patients were asked to name as many words as possible in one minute, starting with a specific letter. All existing nouns were allowed. However, personal names and names of places were not allowed. This part was administered three times, each time using a

different letter, namely D, A and T at the first measurement, K, O and M at the second measurement and P, G and R at the third measurement.[7] For both subtests, the total number of correctly produced words was counted.

To assess working memory, the Digit Span Backward test was used.[3] Patients were presented with sequences of digits, which they had to repeat backwards. The test started with sequences of two digits which expanded until the maximum of eight digits. When a patient failed to correctly repeat two sequences of the same length, the test was aborted. The total number of sequences that was repeated correctly was registered. Different strings of digits were used during each measurement.

The Trail Making Test assesses visual scanning, speed of processing (condition A) and cognitive flexibility (condition B).[8] During condition A, patients were asked to connect 25 circles containing numbers as quickly as possible in ascending order. Part B consisted of 25 circles, containing both numbers and letters. Patients were again required to connect the circles as fast as possible in ascending order, this time alternating between numbers and letters i.e. 1-A-2-B-3-C, etc.. The target measure for cognitive flexibility was the performance on condition B corrected for condition A (Trail Making Test B/A). Parallel versions of the test were performed during each measurement.

Social cognition

Social cognition can be defined as the mental operations underlying social interactions, including perceiving, interpreting, and generating responses to the intentions, dispositions, and behaviors of others.[9] In this study, social cognition was assessed with the Reading the Mind in the Eyes Test (RMET).[10] During this test, patients were presented with 36 photographs, one after each other, of the eye-region of the face. Along with each pair of eyes, four words describing an emotion were presented. Patients were asked to choose which word described best what the person in the picture is feeling or thinking. To exclude any influence of difficulties with word comprehension, a glossary with the presented words was made available.

Reference data: Healthy control subjects

The performances of patients were compared to published normative datasets from healthy populations.

Memory

The results of the RBMT were compared to a healthy reference group (n = 344) consisting of participants aged 17–89.[11] The scores of patients were corrected for age, sex and education. In the study by Schmand and colleagues participants were presented with two stories of comparable length. However, in the present study, patients were presented with only one story. To compensate for this difference, the score obtained

in the present study was doubled. Reference data for the 15 words test were derived from control subjects (n = 847) consisting of participants aged 14–87 for the measures total immediate memory, delayed memory and delayed corrected for total memory.[11] Reference data for the measures short term memory, learning score and recognition of the 15 words test were derived from control subjects of the Maastricht Aging Study.[2] The scores were corrected for age, sex and education. The results of the Digit Span test were compared to the reference group as described in the Wechsler Memory Scale Manual (n = 316). These scores were corrected for age.[3] Patients' scores on the Rey Complex Figure were compared to reference data as provided in the manual of the Rey Complex Figure Test and Recognition Trial and were also corrected for age.[12] No normative data was available for the 15 Figures Test.

Attention

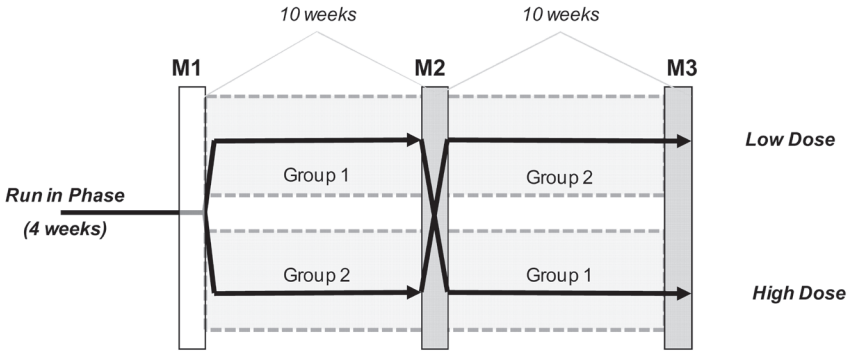
Norm data for the divided attention, visual scanning and alertness subtests were provided by the test authors within the Test of Attentional Performance Version 2.2.[5] A correction for age and sex was applied. No normative data was available for the vigilance subtest.

Executive functioning

The results of the semantic subtest of the Verbal Fluency Test were compared to a reference group (n = 464) consisting of healthy participants aged 17–90.[11] The scores were corrected for age and education. The scores of the phonemic subtest of the Verbal Fluency Test were compared to a reference group (n = 570) consisting of participants aged 17–90 and corrected for education.[11] The results of the TMT were compared to a reference group (n = 478), consisting of healthy participants aged 17–90.[11] The scores were corrected for age, sex and education.

Social cognition

The performance on the RMET was compared to a healthy reference group (n = 200) and corrected for sex.[13]



Supplemental Figure 1. Schematic representation of the study design. M1: Baseline measurement; M2: Measurement after first treatment period; M3: Measurement after second treatment period.

Supplemental Table 1. Cognitive performances of pituitary patients treated for adrenal insufficiency at baseline

	Group 1 (first low dose followed by high dose) (N = 22)	Group 2 (first high dose followed by low dose) (N = 25)	*P*-value*
Memory			
RBMT †			
Immediate memory	-0.80 (1.05)	-0.83 (0.87)	0.949
Delayed memory	-0.69 (0.76)	-0.86 (0.77)	0.535
Delayed corrected for immediate memory	0.08 (1.12)	-0.41 (1.41)	0.311
15 Words test †			
Short-term memory	-0.01 (0.94)	-0.59 (0.79)	0.030[a]
Total immediate memory	-0.24 (1.22)	-0.78 (0.75)	0.060
Learning score	0.17 (1.18)	0.09 (0.85)	0.831
Delayed memory	-0.31 (1.26)	-0.78 (0.80)	0.281
Delayed corrected for total memory	-0.28 (0.94)	-0.30 (0.73)	0.915
Recognition	-0.13 (1.11)	-0.22 (0.65)	0.126
Digit Span forward †			
Short-term memory	0.15 (0.95)	0.47 (1.00)	0.330
Rey Complex Figure ‡			
Immediate memory	0.70 (1.32)[b]	0.64 (1.51)	0.921
Delayed recall	0.64 (1.24)[b]	0.74 (1.43)	0.691
15 Figures Test *(raw data)* ‡[Δ]			
Total memory	42.91 (13.27)	44.04 (12.46)	0.915
Delayed memory	10.55 (3.31)	10.92 (3.00)	0.780
Recognition	26.86 (6.39)	28.68 (1.55)	0.372
Attention			
Vigilance *(raw data)* ‡ [Δ]			
Reaction time of correct responses	606.24 (84.27)[b]	643.16 (131.02)	0.402
Variability of reaction time	128.67 (135.43)[b]	131.44 (62.85)	0.120
Number of omission errors	0.62 (0.87)[b]	1.44 (1.36)	0.463
Number of commission errors	9.52 (37.96)[b]	2.04 (2.65)	0.026[a]
Divided attention ‡			
Reaction time auditory responses	-1.13 (0.88)	-0.92 (0.82)	0.288
Variability of reaction time auditory responses	0.16 (1.06)	0.10 (1.15)	0.473
Reaction time visual responses	-0.48 (0.82)	0.05 (0.97)	0.035
Variability of reaction time visual responses	-0.46 (1.21)	-0.05 (0.96)	0.175
Number of omission errors	-0.37 (0.93)	-0.39 (0.93)	0.974
Number of commission errors	0.13 (1.05)	0.23 (0.97)	0.842

Supplemental Table 1. Cognitive performances of pituitary patients treated for adrenal insufficiency at baseline (continued)

	Group 1 (first low dose followed by high dose) (N = 22)	Group 2 (first high dose followed by low dose) (N = 25)	*P*-value*
Visual scanning ‡			
Reaction time for target stimuli	-0.87 (0.73)	-0.75 (0.90)	0.886
Variability of reaction time for target stimuli	-0.78 (0.74)	-0.72 (0.88)	0.912
Number of omission errors	-0.56 (1.03)	0.34 (1.16)	0.015
Number of commission errors	-0.78 (0.32)	-0.72 (0.36)	0.346
Alertness ‡			
Tonic alertness reaction time	-0.75 (1.29)	-0.80 (0.76)	0.583
Tonic alertness variability in reaction time	-0.08 (0.85)	-0.35 (0.98)	0.410
Phasic alertness reaction time	-1.17 (0.56)	-0.92 (0.77)	0.366
Phasic alertness variability in reaction time	-0.31 (0.65)	-0.19 (0.91)	0.492
Executive functioning			
Fluency test †			
Semantic fluency	-0.21 (1.00)	-0.04 (1.32)	0.789
Phonemic fluency	0.12 (1.39)	-0.34 (0.82)	0.709
Digit Span backwards †			
Working memory	0.29 (0.83)	0.18 (1.12)	0.398
Trail Making Test ‡			
Condition A, time	-0.80 (1.01)	-0.52 (1.21)	0.522
Condition B, time	-0.43 (0.94)	-0.38 (0.97)	0.773
Condition B/A	-0.04 (0.99)	-0.15 (0.84)	0.631
Social cognition			
Reading the Mind in the Eyes Test †‡	-0.85 (1.05)	-0.80 (0.90)	0.923

Cognitive performance data are given as mean Z-scores (SD); RBMT: Rivermead Behavioral Memory Test; * Group 1 versus Group 2 by Mann-Whitney U test;

†: verbal test; ‡: non-verbal test; [a] *P*-value not statistically significant after Bonferroni correction; [b] N = 21; [Δ] Performances on these cognitive tests are given as raw data mean (SD) due to lacking reference data

REFERENCES

1. Wilson BA, Cockburn J, Baddeley AD. The Rivermead Behavioural Memory Test. Bury St Edmunds: Thames Valley Test Company; 1985.
2. Van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J. Rey's verbal learning test: Normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *J Int Neuropsychol Soc* 2005;11:290-302.
3. Wechsler D. Wechsler Memory Scale Revised. The Psychological Corporation; 1987.
4. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. *Archives de Psychologie* 1941;28:286-340.
5. Zimmermann P, Fimm B. Test of Attentional Performance Version 2.2. Psychologische Testsysteme; 2007.
6. Straus E, Sherman E, Spreen O. A compendium of neuropsychological test: administration norms and commentary. New York: University Oxford Press; 2006.
7. Schmand B, Groenink SC, van den Dungen M. Letter fluency: Psychometric properties and dutch normative data. *Tijdschr Gerontol Geriatr* 2008;39:64-76.
8. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept. *Mot Skills* 1958;8:271-276.
9. Davis MC, Lee J, Horan WP, Clarke AD, McGee MR, Green MF, Marder SR. Effects of single dose intranasal oxytocin on social cognition in schizophrenia. *Schizophr Res* 2013;147:393-7.
10. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry* 2001;42:241-251.
11. Schmand B, Houx P, de Koning I. Normen neuropsychologische tests 2012. Nederlands Instituut van Psychologen; 2012.
12. Meyers JE, Meyers KR. RCFT Rey Complex Figure Test and Recognition Trial: Professional Manual. Ltz, FL. Psychological Assessment Recources; 1995.
13. Vellante M, Baron-Cohen S, Melis M, Marrone M, Petretto DR, Masala C, Preti A. The "Reading the Mind in the Eyes" Test: systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry* 2013;18:326-354.